Combining Conformational Flexibility and Continuum Electrostatics for Calculating pK_a's in Proteins

R.E. Georgescu, E.G. Alexov, and M.R. Gunner

Supplementary material: Biophysical Journal Oct. 2002

SM I-X: Experimental and calculated pK_a's for residues with measured pK_a's. Residue name is underlined if it has electrostatic interactions with another residue of more than –3 ΔpK units (–4.08 kcal/mol). If the residue name is in bold the ionized form has lost more than 3 ΔpK units of reaction field energy ($\Delta\Delta G_{rxn}$ > 4.08 kcal/mol). Only the initial heavy atom conformer in the starred protein data file is used to determine if a residue is in either category.

Expt: Experimental pK_a's: ±: Reported error of experimental measurements; n: reported Hill coefficient value (underlined, if the n values were explicitly reported otherwise n was estimated from the shape of published titration curves).

The columns headed by protein data file name contains the calculated pK_a's. Err: Difference between calculated and experimental value. Absolute errors greater than 1 pH unit in bold while those greater than 2 pH units are also underlined. #: SCCE: Single Conformation Continuum Electrostatics. All other calculations obtained with MCCE. *:Representative calculation for this protein. §: Calculations using NMR structures. All other coordinates obtained by x-ray crystallography. For the NMR structures the pK_a is the average value for the ensemble of structures; RMSV: Root-mean-square-variation of pK_a's calculated for a given residue (J) in the group of n structures is $sqrt(\frac{1}{n}\sum_{J=1}^{n}(pK_{av} - pK_J)^2)$. All calculations carried out with 0.15 M ionic strength and ϵ_{prot}=4 except for ¶ where ϵ_{prot}=8 and || where ϵ_{prot}=20. RMS: root-mean-square error of computed pK_a's for a given protein data file is $sqrt(\frac{1}{N}\sum_{i=1}^{N}(pK_{exp} - pK_{calc})^2)$ where N is the number of residues with known pK_a's; Conf: Number of conformers in the calculation. Several proteins have ions that influence residue pK_a's. +: Ion included; –: no ion.

SM I. *Bacillus amyloliquefaciens* barnase.

X-ray structures: 1A2P (Martin et al., 1999) and 1B20 (Buckle et al., 1993) and 20 NMR structures in 1BNR (Bycroft et al., 1991). Experimental pK_a's (Oliveberg et al., 1995). The pK_a of His18 (Loewenthal et al., 1993)

		Expt	±	n	#1A2P	#Err	*1A2P	*Err	1B20	Err	§1BNR	§Err	§RMSV	¶1A2P	¶Err	||1A2P	||Err
Asp	8	3.06	0.1	0.9	1.1	**-2.0**	1.4	**-1.7**	1.7	**-1.4**	2.0	**-1.0**	1.2	3.4	0.3	3.7	0.6
Asp	12	3.65	0.1	0.9	2.6	**-1.1**	3.7	0.0	3.6	-0.1	3.8	0.2	0.8	3.4	-0.2	3.7	0.1
Asp	22	3.30	0.1	0.9	2.5	-0.8	2.7	-0.6	3.7	0.4	3.7	0.4	0.6	3.5	0.2	3.7	0.4
Asp	44	3.50	0.1	0.9	4.1	0.6	4.0	0.5	3.9	0.4	3.9	0.4	0.5	4.3	0.8	4.5	**1.0**
Asp	**54**	2.2	0.3	0.9	-1.1	**-3.3**	2.3	0.1	2.8	0.6	1.8	-0.4	2.6	1.3	-1.0	2.5	0.4
Asp	**75**	3.10	0.2	\	3.9	0.8	4.5	**1.4**	4.8	**1.7**	4.6	**1.5**	2.3	1.4	**-1.8**	1.5	**-1.6**
Asp	**86**	4.20	0.1	0.8	5.2	**1.0**	4.9	0.6	4.8	0.6	5.9	**1.7**	1.8	4.9	0.7	4.1	-0.1
Asp	93	2	0.3	\	-1.1	**-3.1**	1.4	-0.6	2.7	0.7	1.6	-0.4	0.8	2.4	0.4	3.3	**1.3**
Asp	101	2.00	0.2	1.0	-0.9	**-2.9**	3.2	**1.2**	3.3	**1.3**	3.0	1.0	1.3	2.6	0.6	3.5	**1.5**
Ctr	**110**	3.30	0.1	0.9	-0.4	**-3.7**	2.7	-0.7	1.9	**-1.4**	3.4	0.1	1.6	2.0	**-1.3**	1.8	**-1.5**
Glu	29	3.75	0.05	0.8	1.4	**-2.4**	2.6	**-1.2**	2.3	**-1.5**	4.5	0.8	0.8	4.1	0.4	4.3	0.5
Glu	60	3.20	0.1	0.9	4.1	0.9	1.8	**-1.4**	2.9	-0.3	3.6	0.4	1.3	3.1	-0.1	3.4	0.2
Glu	**73**	2.1	0.1	1.2	-1.0	**-3.1**	0.6	**-1.5**	-0.8	**-2.9**	3.6	**1.5**	3.2	0.6	**-1.5**	2.5	0.4
His	18	7.72	0.05	1.0	7.0	-0.7	6.6	**-1.1**	7.0	-0.8	6.8	-0.9	0.1	6.6	**-1.1**	6.6	**-1.2**
RMS						2.11		1.04		1.26		0.94	1.63		0.91		0.89
Conf						96		526		480		291			450		299

SM II. Bovine Pancreatic Trypsin Inhibitor .

X-ray structures: 4PTI (Marquart et al., 1983) and 20 NMR structures in 1PIT (Berndt et al., 1992). Experimental pK_a's (Brown et al., 1976; March et al., 1982)

		Expt.	Err.	n	#4PTI	#Err.	*4PTI	Err	§1PIT	§Err	§RMSV	¶4PTI	¶Err	||4PTI	||Err
Asp	3	3.55	0.1	\	3.8	0.3	3.9	0.4	3.7	0.1	0.3	4.0	0.5	4.3	0.7
Asp	50	3.20	0.0	\	3.4	0.2	2.3	-0.9	2.6	-0.6	1.0	2.9	-0.3	3.0	-0.2
Ctr	58	3.10	0.3	0.9	3.4	0.3	3.7	0.6	2.7	-0.4	0.5	3.8	0.7	3.8	0.7
Glu	7	3.85	0.1	0.8	5.6	**1.7**	2.6	**-1.3**	3.1	-0.7	1.4	4.5	0.6	4.1	0.2
Glu	49	3.91	0.1	\	3.2	-0.7	3.9	0.0	3.5	-0.4	1.1	4.4	0.5	4.4	0.5
Lys	15	10.43	0.1	\	10.9	0.5	10.9	0.5	10.4	-0.1	0.4	11.0	0.5	10.8	0.4
Lys	26	10.44	0.1	\	10.5	0.1	10.6	0.2	10.6	0.1	0.4	10.8	0.3	10.7	0.3
Lys	41	10.75	0.1	\	12.1	**1.4**	11.6	0.8	11.2	0.5	1.0	11.4	0.6	11.3	0.5
Lys	46	10.35	0.2	\	10.6	0.2	10.3	-0.0	10.2	-0.2	0.6	10.4	0.1	10.3	-0.0
Ntr	1	7.90	0.2	1.0	6.5	**-1.4**	7.7	-0.2	8.3	0.4	0.8	7.8	-0.1	7.9	-0.0
Tyr	10	9.46	0.0	0.7	9.0	-0.4	9.4	-0.0	9.7	0.3	1.6	9.4	-0.1	9.9	0.4
Tyr	21	9.94	0.0	0.8	10.1	0.2	10.3	0.4	10.6	0.6	0.7	10.2	0.3	10.4	0.5
Tyr	**23**	11	0.1	\	12.0	1.0	12.1	**1.1**	10.6	-0.4	1.5	11.0	0.0	10.8	-0.2
Tyr	**35**	10.60	0.1	0.8	10.5	-0.1	10.0	-0.6	12.4	**1.8**	3.9	9.7	-0.9	10.4	-0.2
RMS						0.80		0.64		0.62	1.40		0.47		0.41
Conf						53		187		140			138		131

SM-III. Intestinal Bovine Calcium-Binding Protein (CabD).

X-ray structure: 3ICB (Szebenyi and Moffat, 1986) and the 24 NMR structures in 1CDN (Akke et al., 1995) have calcium bound (holo-enzyme). The 33 NMR structures in 1CLB (Skelton et al., 1995) do not (apo-form). Experimental pK_a's (Kesvatera et al., 1996; 1999). The reported error of experimental pK_a values is less than 0.08.

Ca^{+2}		pK_a +	pK_a –	n	3ICB +	Err	*3ICB +	Err	§1CDN +	§Err	§RMSV	§1CLB –	§Err	§RMSV	¶3ICB +	¶Err	‖3ICB +	‖Err
Asp	47		3.00	0.7								2.9	-0.1	0.7				
Ctr	75		3.20	1.1								4.6	**1.4**	1.8				
Glu	4		3.80	0.8								3.5	-0.3	0.7				
Glu	5		3.40	0.8								3.1	-0.3	1.2				
Glu	11		4.70	0.7								4.9	0.2	0.8				
Glu	17		3.62	0.6								4.1	0.5	2.9				
Glu	26		4.10	0.9								2.7	**-1.4**	1.8				
Glu	48		4.60	0.6								4.4	-0.2	0.9				
Glu	64		3.80	0.8								4.3	0.5	0.7				
Lys	1	10.60	10.60	0.7	11.0	0.4	10.8	0.2	11.3	0.7	1.3	9.7	-0.9	2.2	11.5	0.9	11.8	**1.2**
Lys	7	11.35	11.39	0.9	10.9	-0.4	10.7	-0.7	11.5	0.2	1.0	10.9	-0.5	1.0	11.4	0.1	11.8	0.5
Lys	12	10.82	11.06	0.7	10.9	0.1	10.8	-0.0	11.2	0.3	0.5	10.9	-0.1	0.5	11.3	0.4	11.8	0.9
Lys	16	10.09	10.07	0.7	11.1	**1.0**	10.6	0.5	10.1	0.0	2.2	11.2	**1.1**	2.2	11.0	0.9	11.5	**1.4**
Lys	25	11.81	11.69	0.9	13.3	**1.4**	13.4	**1.5**	12.2	0.4	0.9	11.5	-0.1	0.7	12.7	0.9	12.7	0.9
Lys	29	10.97	11.37	0.8	9.5	**-1.5**	10.4	-0.5	10.5	-0.4	1.3	10.2	**-1.2**	1.6	11.0	0.0	11.3	0.3
Lys	41	10.89	10.93	0.8	10.9	0.0	10.9	0.0	10.8	-0.1	0.9	10.6	-0.3	0.9	10.9	0.0	11.0	0.1
Lys	**55**	11.38	12.12	0.9	10.2	**-1.2**	11.6	0.2	12.4	1.0	1.1	11.7	-0.4	0.8	11.3	-0.0	12.2	0.9
Lys	71	10.72	10.73	0.8	10.1	-0.6	9.9	-0.8	11.1	0.4	0.7	10.2	-0.5	1.3	10.4	-0.3	10.7	0.0
Lys	72	10.96	11.33	0.8	10.8	-0.2	10.5	-0.5	11.2	0.2	0.9	10.7	-0.7	1.3	10.9	-0.0	11.3	0.3
RMS						0.87		0.66		0.47	1.17		0.71	1.27		0.38		0.59
Conf						72		273		336			237			175		146

SM-VI. Rat T-Lymphocyte Adhesion Glycoprotein (CD2).

X-ray structure 1HNG (Jones et al., 1992)

		Expt	±	n	#1HNG	#Err	*1HNG	*Err	¶1HNG	¶Err	‖1HNG	‖Err
Asp	2	3.50	0.02	0.9	3.1	-0.4	3.0	-0.5	3.4	-0.1	3.8	0.3
Asp	25	3.53	0.02	0.9	3.8	0.3	3.9	0.4	4.2	0.6	4.4	0.8
Asp	26	3.58	0.03	0.9	3.9	0.4	4.1	0.5	4.2	0.7	4.5	0.9
Asp	**28**	3.57	0.06	0.8	1.5	**-2.1**	1.9	**-1.7**	2.8	-0.8	3.0	-0.6
Asp	**62**	4.18	0.03	0.8	6.5	**2.4**	5.8	**1.6**	5.2	**1.0**	4.8	0.6
Asp	71	3.20	0.10	0.8	3.8	0.6	3.8	0.6	3.9	0.7	4.4	**1.2**
Asp	**72**	4.14	0.05	0.8	4.4	0.2	4.8	0.7	5.0	0.9	3.9	-0.2
Asp	94	3.83	0.04	0.8	3.0	-0.8	3.4	-0.4	3.9	0.1	3.6	-0.2
Ctr	99	3.11	0.05	0.8	-1.0	**-4.1**	3.8	0.7	3.8	0.7	3.9	0.8
Glu	29	4.51	0.10	0.6	1.8	**-2.7**	4.8	0.3	3.0	**-1.5**	3.7	-0.9
Glu	33	4.20	0.04	0.8	4.7	0.5	4.5	0.3	4.5	0.3	4.3	0.1
Glu	**41**	6.53	0.10	0.5	7.5	0.9	5.8	-0.7	5.3	**-1.2**	5.0	**-1.5**
Glu	56	3.95	0.02	0.8	3.9	-0.0	4.0	0.0	4.0	0.1	3.9	-0.0
Glu	99	4.10	0.05	0.8	4.6	0.5	4.0	-0.1	4.0	-0.1	3.9	-0.2
RMS						1.62		0.77		0.76		0.73
Conf						154		398		309		285

SM-V. Hen Egg-White Lysozyme (HEWL).

X-ray structures: 1B0D (Vaney et al., 1996); 2LZT (Ramanadham et al., 1981), 1HEL (Wilson et al., 1992). NMR structure 1E8L (Schwalbe et al., 2001). Experimental data (Kuramitsu and Hamaguchi, 1980; Bartik et al., 1994). SCCE columns show only errors of calculated pK_a's for 2LZT and 1B0D.

		Expt	±	n	#2LZT err	#1BOD err.	*1BOD	Err	2LZT	Err	2HEL	Err	§1E8L	§Err	§RMSV	¶2LZT	¶Err	||2LZT	||Err
Asp	18	2.66	0.1	0.8	-0.1	0.3	3.0	0.4	2.8	0.2	2.2	-0.5	3.6	0.9	0.5	3.6	1.0	3.5	0.8
Asp	**48**	1.60	0.4	\	-0.7	**2.1**	0.6	**-1.0**	1.7	0.1	-0.2	**-1.8**	3.2	**1.6**	0.2	2.6	1.0	4.1	**2.5**
Asp	**52**	3.68	0.1	0.8	**-3.4**	**-4.1**	3.9	0.2	3.0	-0.7	4.6	0.9	3.8	0.1	2.5	4.5	0.9	3.5	-0.2
Asp	**66**	0.90	0.5	\	**1.9**	**1.9**	1.5	0.6	2.6	**1.7**	1.8	0.9	2.7	**1.8**	0.8	1.4	0.5	1.9	**1.0**
Asp	87	2.07	0.2	\	0.7	**3.1**	1.0	**-1.1**	1.2	-0.9	1.2	-0.9	1.7	-0.3	0.3	2.7	0.6	3.5	**1.5**
Asp	101	4.08	0.1	\	**-4.7**	0.5	3.8	-0.3	5.1	1.0	4.3	0.2	4.1	0.0	0.2	4.2	0.1	4.1	0.1
Asp	119	3.20	0.1	\	-0.1	-0.3	4.1	0.9	2.4	-0.8	3.8	0.6	3.3	0.1	0.4	4.4	**1.2**	3.7	0.5
Ctr	129	2.75	0.1	\	**-3.8**	0.8	2.4	-0.4	2.6	-0.1	3.0	0.2	2.5	-0.3	0.6	2.5	-0.2	3.0	0.2
Glu	**7**	2.85	0.3	\	1.0	**2.4**	3.5	0.6	2.2	-0.6	2.2	-0.6	3.3	0.5	0.2	3.1	0.2	3.5	0.7
Glu	**35**	6.20	0.1	\	0.8	**1.3**	6.2	0.0	6.2	-0.0	6.6	0.4	6.8	0.6	6.8	5.7	-0.5	4.8	**-1.4**
His	15	5.36	0.1	\	**-1.5**	**-1.8**	6.5	**1.1**	6.4	**1.0**	6.8	**1.5**	5.8	0.5	0.3	5.7	0.3	5.5	0.1
Lys	1	10.80	0.1	\	0.8	**-1.7**	10.3	-0.5	9.8	-1.0	10.2	-0.6	10.3	-0.5	1.0	12.1	**1.3**	10.3	-0.5
Lys	13	10.50	0.1	\	**3.8**	-0.9	11.2	0.7	10.8	0.3	10.7	0.2	10.9	0.4	0.4	11.2	0.7	11.2	0.7
Lys	33	10.40	0.1	\	**1.9**	**1.3**	9.8	-0.6	8.9	**-1.5**	9.5	-0.9	9.1	**-1.3**	1.4	10.3	-0.1	10.2	-0.2
Lys	96	10.80	0.1	\	0.7	**-1.4**	11.3	0.5	11.9	**1.1**	11.8	1.0	10.8	0.0	0.8	12.2	1.4	11.5	0.7
Lys	97	10.30	0.1	\	1.0	-0.6	10.5	0.2	11.1	0.8	10.8	0.5	10.6	0.3	0.7	10.8	0.5	11.4	**1.1**
Lys	116	10.20	0.1	\	0.2	**1.5**	10.0	-0.2	9.5	-0.7	9.4	-0.8	9.8	-0.4	0.7	10.5	0.3	10.4	0.2
Tyr	20	10.30	0.1	\	**-1.8**	**1.3**	9.7	-0.6	11.1	0.8	9.6	-0.7	9.7	-0.6	1.0	9.5	-0.8	10.4	0.1
Tyr	23	9.80	0.1	\	-0.8	0.5	9.3	-0.5	9.3	-0.5	9.7	-0.1	9.3	-0.5	0.7	9.6	-0.2	9.9	0.1
RMS					2.05	1.74		0.63		0.86		0.81		0.73	1.77		0.74		0.90
Conf						104		451		466		443		253			331		288

SM- VI. Third domain of the turkey ovomucoid inhibitor (OMTKY3).

X-ray structures: 1PPF (Bode et al., 1986), 1CHO (Fujinaga et al., 1987) and 3SGB (Read et al., 1983). The 50 NMR structures in 1OMU (Hoogstraten et al., 1995). Experimental data (Schaller and Roberston, 1995; Forsyth et al., 1998)

		Expt	±	n	#1PPF	#Err	*1PPF	Err	1CHO	Err	3SGB	Err	§1OMU	§Err	§RMSV	¶1PPF	¶Err	‖1PPF	‖Err
Asp	7	**2.7**	0.1	0.7	0.4	**-2.3**	2.8	0.1	4.2	**1.5**	2.7	-0.0	2.9	0.3	0.5	4.0	**1.3**	5.0	**2.3**
Asp	**27**	**2.3**	0.1	0.9	3.9	**1.6**	3.3	0.9	2.7	0.4	2.3	0.0	3.2	0.9	0.5	4.0	**1.7**	3.6	**1.3**
Ctr	56	2.2	0.2	0.9	2.7	0.5	3.2	0.9	3.2	0.9	3.2	0.9	3.3	**1.1**	0.9	3.2	1.0	3.3	**1.1**
Glu	10	4.1	0.2	0.9	3.7	-0.4	3.5	-0.6	3.9	-0.2	4.2	0.1	3.7	-0.4	0.6	3.7	-0.4	3.8	-0.3
Glu	**19**	3.2	0.0	1.1	-0.3	**-3.5**	1.6	**-1.6**	2.9	-0.3	2.7	-0.5	2.7	-0.5	0.6	3.9	0.7	4.3	**1.1**
Glu	43	4.8	0.2	1.0	4.6	-0.2	4.5	-0.3	4.6	-0.2	4.6	-0.2	4.8	-0.0	0.1	4.6	-0.2	4.9	0.1
His	52	7.5	\	0.9	8.5	1.0	8.2	0.7	8.6	**1.1**	8.5	1.0	7.1	-0.4	0.7	7.5	-0.0	6.9	-0.6
Lys	13	9.9	0.0	0.7	14.3	**4.4**	11.5	**1.6**	10.1	0.2	10.0	0.1	11.2	**1.3**	1.6	11.4	**1.5**	12.0	**2.1**
Lys	29	11.1	0.1	0.9	11.7	0.6	11.2	0.1	10.9	-0.2	10.8	-0.3	10.4	-0.7	1.1	11.4	0.3	11.3	0.2
Lys	34	10.1	0.1	0.7	12.4	**2.3**	7.2	**-2.9**	8.8	**-1.3**	10.1	0.0	7.6	**-2.5**	1.3	12.6	**2.5**	12.0	**1.9**
Lys	55	11.1	0.1	0.6	10.3	-0.8	10.2	-0.9	10.8	-0.3	11.7	0.6	10.8	-0.3	0.6	10.7	-0.4	10.9	-0.2
Ntr	1	8.0	0.0	0.9	5.9	**-2.0**	7.3	-0.7	7.3	-0.7			8.4	0.4	0.5	7.4	-0.5	7.6	-0.4
Tyr	**11**	10.2	0.1	0.7	8.9	**-1.3**	11.8	**1.6**	13.7	**3.5**	14.6	**4.4**	12.7	**2.5**	1.5	10.3	0.1	10.5	0.3
Tyr	20	11.1	0.2	0.6	9.8	**-1.3**	10.1	-1.0	10.8	-0.3	9.4	**-1.7**	10.6	-0.5	0.7	9.8	**-1.3**	10.1	-1.0
Tyr	**31**	**12.5**	\	\	15.0	**3.0**	12.2	-0.3	14.4	**1.9**	13.0	0.5	12.5	0.0	0.3	11.4	**-1.1**	13.9	**1.5**
RMS						2.06		1.19		1.24		1.34		1.10	0.87		1.10		1.19
Conf						51		213		219		167		164			154		122

SM-VII. B1 immunoglobulin-binding domain of streptococcal protein G.

X-ray structures 1PGA (Gallagher et al., 1994), 60 NMR structures in 1GBL (Gronenborn et al., 1991). Experimental pK_a's (Khare et al., 1997). Estimated n-values are all in the range 0.9 - 1.

		Expt	±	#1PGA	#Err	*1PGA	*Err	§1GBL	§Err	§RMSV	1PGA	Err	‖1PGA	‖Err
Asp	22	2.9	0.1	0.5	**-2.4**	2.2	-0.7	2.6	-0.3	1.0	3.1	0.2	3.7	0.8
Asp	36	3.8	0.1	5.3	**1.5**	5.0	**1.2**	4.7	0.9	0.7	5.2	**1.4**	5.0	**1.2**
Asp	40	4.0	0.2	4.5	0.5	4.1	0.1	5.4	**1.4**	0.9	4.7	0.7	5.0	**1.0**
Asp	46	3.6	0.1	1.6	**-2.0**	3.1	-0.5	3.7	0.1	0.3	4.2	0.6	4.4	0.8
Asp	47	3.4	0.3	0.3	**-3.2**	2.3	**-1.1**	4.4	**1.0**	0.2	3.0	-0.4	3.2	-0.2
Ctr	56	4.0	0.1	3.0	-1.0	3.3	-0.7	4.3	0.3	0.9	3.5	-0.5	3.7	-0.4
Glu	**56**	4.0	0.1	3.9	-0.1	5.3	**1.3**	6.8	**2.8**	0.8	4.9	0.9	4.7	0.7
Glu	15	4.4	0.1	1.7	**-2.7**	3.8	-0.6	4.3	-0.1	0.6	3.8	-0.6	4.1	-0.3
Glu	19	3.7	0.2	2.7	-1.0	3.5	-0.2	4.2	0.5	0.2	3.9	0.2	4.0	0.3
Glu	27	4.5	0.1	-0.5	**-5.0**	3.8	-0.7	4.5	0.0	0.2	3.9	-0.6	3.3	**-1.2**
Glu	42	4.4	0.1	5.0	0.6	5.1	0.7	5.0	0.6	0.5	5.0	0.6	5.2	0.8
Lys	10	11.0	\	11.9	0.9	11.4	0.4	11.1	0.1	0.4	11.2	0.2	11.5	0.5
Lys	13	11.0	\	11.0	-0.0	10.8	-0.2	10.8	-0.2	0.6	11.0	-0.0	11.5	0.5
Lys	**28**	10.9	\	13.4	**2.5**	11.7	0.8	10.2	-0.7	0.7	11.8	0.9	12.0	**1.1**
Tyr	**33**	11.0	\	10.7	-0.3	10.8	-0.2	10.3	-0.7	0.2	10.5	-0.5	10.9	-0.1
RMS					2.13		0.67		0.65	0.60		0.63		0.75
Conf					55		240		202			143		107

SM-VIII Ribonuclease A (RNase A).

X-ray structures 3RN3 (Howlin et al., 1989), 7RSA (Wlodawer et al., 1988) and the 32 NMR structures in 2AAS (Santoro et al., 1993). Experimental values (Antosiewicz et al., 1996). His pK_a values averages experimental data (Ruterjans and Witzel, 1969; Matthews and Westmoreland, 1973; Walters and Allerhand, 1980). Only the pK_a's of His 12 and 19 are effected by added PO_4. +: with PO_4. Experimental values from (Meadows et al., 1969; Cohen et al., 1973).

Calculations had PO_4 dynamically equilibrated with the protein during Monte Carlo sampling. –: no PO_4.

		Expt	±	#3RN3	#Err.	*3RN3	*Err.	7RSA	Err.	§2AAS	§Err	§RMSV	¶3RN3	¶Err.	3RN3	Err.
PO_4				+		+		+		–			+		+	
Asp	**14**	2	0.2	-1.0	**-3.0**	0.6	**-1.4**	0.6	**-1.4**	4.9	**2.9**	5.5	1.3	-0.7	2.8	0.8
Asp	38	3.1	0.2	3.0	-0.1	3.0	-0.1	1.7	**-1.4**	3.1	0.0	0.4	3.6	0.5	3.8	0.7
Asp	53	3.9	0.2	3.4	-0.5	3.7	-0.2	3.4	-0.5	4.0	0.1	1.2	4.3	0.4	4.3	0.4
Asp	**83**	3.5	0.2	6.3	**2.8**	3.8	0.3	3.2	-0.3	2.7	-0.8	1.2	3.6	0.1	3.1	-0.4
Asp	**121**	3.1	0.2	-1.0	**-4.1**	3.2	0.1	0.5	**-2.6**	2.2	-0.9	1.4	2.7	-0.4	3.0	-0.1
Ctr	124	2.4	0.2	-0.0	**-2.4**	1.3	**-1.1**	0.9	**-1.5**	1.6	-0.8	1.0	1.7	-0.7	2.2	-0.2
Glu	2	2.8	0.2	-0.6	**-3.4**	1.3	**-1.5**	5.6	**2.8**	2.4	-0.4	1.5	2.8	0.0	3.2	0.4
Glu	9	4	0.2	5.8	**1.8**	5.4	**1.4**	5.5	**1.5**	5.1	**1.1**	1.2	5.0	1.0	4.6	0.6
Glu	49	4.7	0.2	5.7	1.0	5.5	0.8	5.7	1.0	7.5	**2.8**	2.9	5.3	0.6	5.0	0.3
Glu	86	4.1	0.2	4.4	0.3	4.2	0.1	5.7	**1.6**	4.6	0.5	0.9	4.8	0.7	4.5	0.4
Glu	111	3.5	0.2	4.4	0.9	3.9	0.4	4.3	0.8	4.8	**1.3**	1.6	4.1	0.6	4.2	0.7
His	**12-**	5.8	0.05	-1.0	**-6.8**	4.2	**-1.6**	4.2	**-1.6**	4.7	**-1.1**	2.0				
	12+	7				6.7	-0.3						6.7	-0.3	6.9	-0.1
His	**48**	6.3	0.1	3.1	**-3.2**	8.8	**2.5**	7.5	**1.2**	4.7	**-1.6**	4.4	8.0	**1.7**	6.7	0.4
His	105	6.6	0.15	6.1	-0.5	5.9	-0.7	6.1	-0.5	6.2	-0.4	0.4	6.4	-0.2	6.3	-0.3
His	119-	6.1	0.1	6.7	0.6	5.4	-0.7	6.1	0.0	5.4	-0.7	1.2				
	119+	7.1				6.0	**-1.1**						6.7	-0.4	6.7	-0.3
Ntr	1	7.6	0.15	6.0	**-1.6**	7.4	-0.2	8.7	**1.1**	8.0	0.4	0.7	7.5	-0.1	7.5	-0.1
RMS					2.69		0.98		1.44		1.28	2.21		0.66		0.44
Conf					123		406		422		356			350		285

SM IX. Ribonuclease Hi (RNase H).

X-ray structure: 2RN2 (Katayanagi et al., 1992), 1RNH (Yang et al., 1990), and 1RDD (Katayanagi et al., 1993) and 8 NMR structures in 1RCH (Yamazaki et al., 1997). The experimental pK_a's (in presence or absence of Mg^{+2}) are from (Oda et al., 1993; Oda et al., 1994). Experimental values that differ by more than 0.2 pH units are bold; No pK_a's with Mg^{+2} are reported for His. The reported errors assume that these pK_a's are independent of added ion. *The pK_a's of Asp 10 are not included in the RMS calculations*. 1RNH does not contain coordinates for residues 1-3 and 153 through the C-terminal 155.

Mg^{+2}		Expt −	Expt +	±	#2RN2 −	#Err.	*2RN2 −	*Err.	1RNH −	Err	1RDD +	Err.	§1rch −	§Err	§RMSV	¶2RN2 −	¶Err.	‖2RN2 −	‖Err.
Asp	*10*	*6.1*	*4.2*		*10.9*	***4.8***	*10.4*	***4.3***	*15.4*	***9.3***	*8.7*	***4.5***	*14*	***7.9***	*3.1*	*8.7*	***2.6***	*6.4*	*0.3*
Asp	70	**2.6**	**3.4**	0.6	2.4	-0.2	2.3	-0.3	3.2	0.6	2.4	-1.0	4.0	**1.4**	1.1	3.6	1.0	3.9	**1.3**
Asp	94	3.2	3.3	1	11.9	**8.7**	3.8	0.6	3.4	0.2	1.4	**-1.9**	3.6	0.4	0.6	3.6	0.4	3.7	0.5
Asp	**102**	2.0	2.0	\	-1.0	**-3.0**	2.0	-0.0	1.4	-0.6	1.6	-0.4	2.3	0.3	0.5	1.8	-0.2	-0.5	**-2.5**
Asp	108	**3.2**	**3.5**	0.9	1.7	**-1.5**	2.0	**-1.2**	2.0	**-1.2**	4.8	**1.3**	4.4	**1.2**	1.0	3.1	-0.1	3.0	-0.2
Asp	134	4.1	4.2	0.8	1.0	**-3.1**	2.2	**-1.9**	4.4	0.3	3.8	-0.4	4.2	0.1	0.5	2.7	**-1.4**	3.3	-0.8
Asp	**148**	2.0	2.0	\	6.9	**4.9**	1.1	-0.9	3.9	**1.9**	-1.0	**-3.0**	4.4	**2.3**	3.6	-0.5	**-2.5**	1.3	-0.7
Ctr	155	3.4	3.5	0.8	3.7	0.3	2.7	-0.7			2.2	**-1.3**	2.4	-1.0	1.7	3.7	0.3	3.4	0.0
Glu	6	4.5	4.5	0.9	4.1	-0.4	6.3	**1.8**	4.3	-0.2	5.7	**1.2**	5.4	0.9	1.1	5.1	0.6	3.9	-0.6
Glu	32	3.6	3.6	0.73	1.3	**-2.3**	2.1	**-1.5**	1.6	**-2.0**	1.5	**-2.1**	3.3	-0.3	0.5	2.5	**-1.1**	3.5	-0.1
Glu	**48**	4.4	4.4	0.75	5.3	0.9	5.5	**1.1**	6.4	**2.0**	5.2	0.8	5.1	0.7	2.9	5.6	**1.2**	4.7	0.3
Glu	57	**3.2**	**3.4**	1	1.7	**-1.5**	2.5	-0.8	2.0	**-1.2**	2.6	-0.8	4.5	**1.3**	1.5	3.0	-0.2	3.3	0.1
Glu	61	3.9	4.0	0.9	1.7	**-2.2**	2.9	**-1.0**	3.5	-0.4	0.5	**-3.5**	3.5	-0.4	0.7	3.3	-0.6	3.2	-0.7
Glu	64	4.4	4.5	0.9	4.6	0.2	4.3	-0.1	2.9	**-1.5**	4.8	0.3	5.3	0.9	1.7	4.4	-0.0	4.3	-0.1
Glu	**119**	4.1	4.3	0.75	-0.4	**-4.5**	3.1	**-1.0**	4.4	0.3	5.8	**1.5**	5.3	**1.2**	1.4	3.3	-0.8	3.9	-0.2
Glu	**129**	**3.6**	**4.0**	0.7	0.0	**-3.6**	2.8	-0.8	2.9	-0.7	2.5	**-1.5**	2.2	**-1.4**	1.7	3.2	-0.4	3.0	-0.6
Glu	131	4.3	4.4	0.85	5.1	0.8	5.0	0.7	3.4	-0.9	5.5	**1.1**	5.3	**1.0**	1.2	5.0	0.7	4.8	0.5
Glu	135	4.3	4.3	0.9	4.3	-0.0	4.3	0.0	3.5	-0.8	4.4	0.1	4.3	-0.0	0.3	4.6	0.3	4.6	0.3
Glu	147	4.2	4.2	1	4.3	0.1	4.4	0.2	4.9	0.7	4.5	0.3	4.2	-0.0	0.5	4.7	0.5	4.7	0.5
Glu	154	4.4	4.4	1	2.6	**-1.8**	3.8	-0.6			6.7	**2.3**	4.8	0.4	0.6	4.0	-0.4	4.6	0.2
His	62	7.0		0.9	5.7	**-1.3**	6.7	-0.3	6.5	-0.5	6.4	-0.6	6.1	-0.9	1.0	6.6	-0.4	6.5	-0.5
His	83	5.5		1	5.4	-0.1	5.4	-0.1	4.8	-0.7	5.8	0.3	5.4	-0.1	0.4	5.5	0.0	5.6	0.1
His	114	5.0		0.9	4.2	-0.8	4.5	-0.6	4.2	-0.8	5.3	0.3	2.8	**-2.2**	3.1	4.8	-0.2	5.7	0.7
His	124	7.1		0.85	4.5	**-2.6**	4.5	**-2.6**	8.7	**1.6**	4.9	**-2.2**	4.7	**-2.4**	2.5	5.2	**-1.9**	5.7	**-1.4**
His	127	7.9		0.9	8.1	0.2	7.0	-0.9	8.7	0.8	8.8	0.9	6.5	**-1.4**	1.7	7.3	-0.6	7.7	-0.2
RMS						2.75		1.04		1.06		1.49		1.15	1.60		0.87		0.77
Conf						133		512				456		365			396		268

SM-X. Ribonuclease T1 (RNase T).

X-ray structure: 3RNT (Kostrewa et al., 1989). Experimental values (Inagaki et al., 1981; McNutt et al., 1990).

		Expt	±	n	#3RNT	#Err.	*3RNT	*Err.	¶3RNT	¶Err.	‖3RTN	‖Err.
His	27	7.3	0.1	\	12.0	**4.7**	8.0	0.7	7.6	0.3	7.8	0.5
His	40	7.9	0.2	\	13.0	**5.1**	9.0	**1.1**	8.4	0.5	7.2	-0.7
His	92	7.8	0.2	\	8.0	0.2	7.1	-0.7	6.9	-0.9	7.5	**-0.3**
Glu	**58**	4.3	0.2	\	-0.4	**-4.7**	4.5	0.2	4.2	-0.1	5.2	0.9
RMS						4.18		0.75		0.54		0.63
Conf						95		395		316		244

SM XI. Lennard-Jones parameters (kcal/mol).

$$\Delta G_{ij}^{non.el} = \left[\frac{A}{r^{12}} - \frac{B}{r^{6}} \right].$$

Value varies for atom types. not-H* are heavy atoms other than H, N, C, or O.

Å where $\Delta G^{ij}=0$ is the distance between atom centers where $\Delta G_{ij}^{non.el}=0$,.

E_{min} is the energy minimum for this atom pair is.

No Lennard-Jones interactions are determined for atoms in the same side chain or between CA (in the backbone) and CB (in the side chain) of the same residue.

atoms	A	B	E_{min}	Å where $\Delta G^{ij}=0$
H : H	$1.33x10^{+1}$	$1.81X10^{+0}$	-0.05	1.40
H : C	$1.61X10^{+3}$	$6.66X10^{+1}$	-0.54	1.70
H : O	$4.02X10^{+2}$	$3.33X10^{+1}$	-0.54	1.51
H : N	$6.27X10^{+2}$	$3.73X10^{+1}$	-0.44	1.60
H : not-H*	$1.61X10^{+3}$	$6.66X10^{+1}$	-0.54	1.70
not-H : not-H	$2.96X10^{+5}$	$1.05X10^{+2}$	-0.01	3.76

SM XII. Reaction field energies (kcal/mol) for isolated side chains.

Energy for transfer from a medium with a dielectric constant of ε_{prot} of 4, 8, or 20 to a solvent with ε=80. $\Delta\Delta G_{rxn,i}$ (in eqn. 3) is the difference between this reference value and the reaction field energy for a conformer in the protein. $\Delta\Delta G_{rxn,i}$ is constrained to have a minimum of 0.

Residue	ε= 4.0 ionized	ε= 4.0 neutral	ε= 8.0 ionized	ε=.0 neutral	ε= 20.0 ionized	ε= 20.0 neutral
ASP	18.1	2.1	8.7	1.1	2.9	0.3
CTR	16.6	1.5	7.9	0.8	2.6	0.2
GLU	18.7	2.2	8.9	1.1	2.9	0.4
HIS	11.8	1.8	5.7	0.9	2.1	0.3
LYS	19.6	1.1	9.8	0.6	3.2	0.2
TYR	15.2	1.2	7.3	1.1	2.4	0.3
NTR	13.3	1.8	6.7	0.9	2.1	0.3
ARG	12.8	3.4	8.0	1.7	2.9	0.5
ASN		5.9		1.7		0.5
GLN		5.9		1.7		0.5
SER		1.8		0.8		0.3
THR		1.8		0.8		0.2
HOH		3.6		2.4		1.1
PO_4^{-2}	62.1		32.5		14.8	
Mg^{+2}	17.8		8.9		4.7	